

Méthodes de Quasi Monte-Carlo pour l'évaluation de stratégies d'investissements

J. DEMGNE ^{1 2}
W. LAIR ¹, J. LONCHAMPT ¹, M. BAUDIN ¹
S. MERCIER ²

¹EDF R & D - 6 Quai Watier, Chatou

²Université de Pau et des Pays de l'Adour - Pau

11 Mars 2014

Plan

Introduction

Processus Markovien Déterministe par Morceaux

Analyse de la moyenne et de la distribution de la $VAN(t)$

Méthodes de quantification de la $VAN(t)$

Résultats numériques

Conclusion & Perspectives

- ▶ Investissements sur les actions de maintenance des actifs de production pour garantir la sûreté et créer de la valeur ajoutée pour assurer la compétitivité des actifs de production

Enjeu

Enjeu

Comparaison de stratégies de maintenance pour aider à la prise de décision sur les investissements

Définition

Une stratégie d'investissements définit des plans de remplacement ou de réparation de composants et prévoit leur approvisionnement

Indicateur de comparaison

Valeur Actuelle Nette (VAN) : différence entre les coûts actualisés induits par la stratégie de référence et la stratégie à évaluer

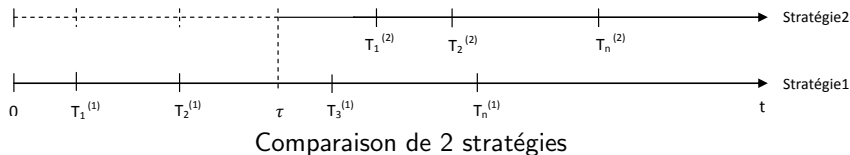
Par exemple :

Stratégie de référence : purement correctif

Stratégie à évaluer : remplacement des composants du parc

Définition de la VAN (t)

$$VAN(t) = C_t^{(1)} - C_t^{(2)}$$

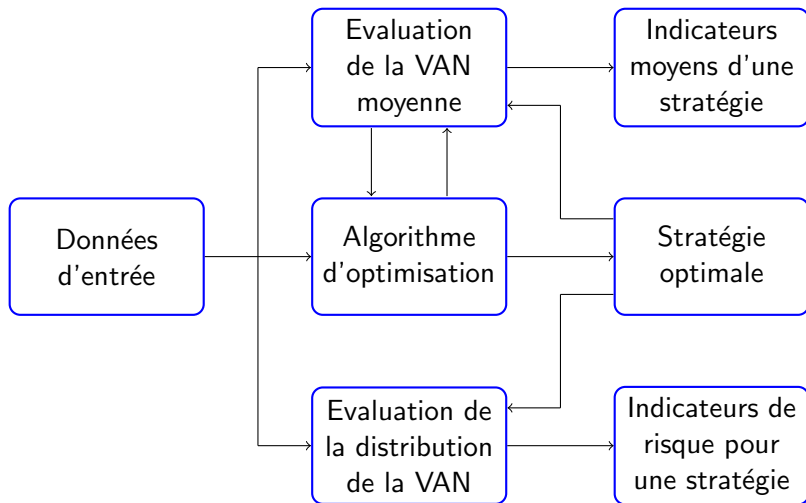


Exemple : τ est la date de remplacement dans la stratégie à évaluer

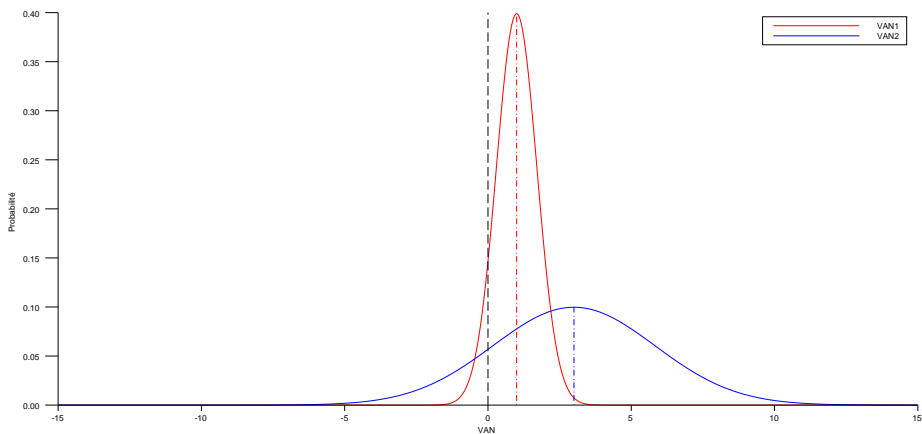
Historique commun entre les deux stratégies

Hypothèse : indépendance des stratégies sachant ce qui s'est passé à l'instant τ

Contexte actuel

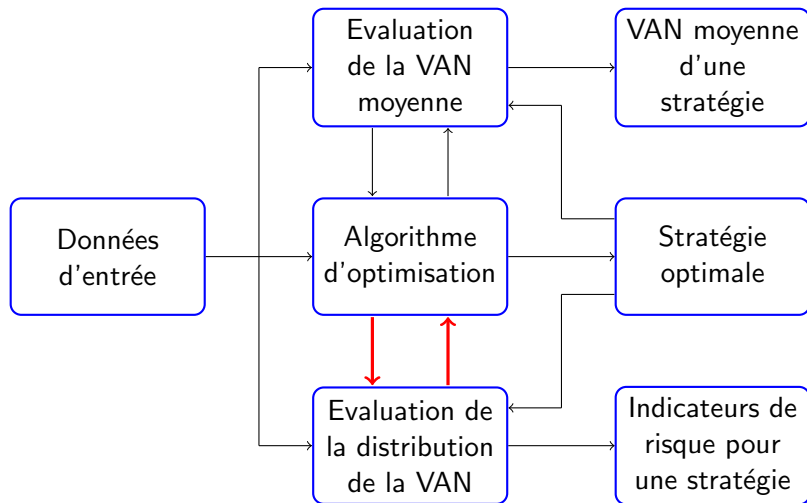


Comparaison de deux stratégies via la loi de $VAN(t)$



Comparaison de 2 distributions de $VAN(t)$

Objectif



Plan

Processus Markovien Déterministe par Morceaux

Définition

Application à notre cadre d'étude

Processus Markovien Déterministe par Morceaux

Processus utilisé en fiabilité dynamique pour modéliser les composants en interaction avec leur environnement

Processus hybride : $(I_t, X_t)_{t \geq 0}$

- ▶ I_t : état des composants à l'instant t (discret)
 - ▶ exemple : marche, marche dégradée, panne
- ▶ X_t : condition environnementale à l'instant t
 - ▶ déterministe entre deux sauts du processus I_t
 - ▶ exemple : température, pression, âge des composants, dates des futures pannes des composants

PDMP - application à notre cadre d'étude

Pour chaque stratégie, nous devons suivre à chaque instant t :

- ▶ l'état des composants du parc, I_t
- ▶ l'état du stock, S_t
- ▶ les dates de futures pannes des composants, X_t
- ▶ le coût cumulé actualisé, C_t

D'où le PDMP décrivant chaque stratégie est de la forme :

$$(Z_t)_{t \geq 0} = (I_t, S_t, X_t, C_t, t)_{t \geq 0}$$

et la chaîne de Markov sous-jacente est :

$$(Z_n)_{n \geq 0} = (I_{T_n}, S_{T_n}, X_{T_n}, C_{T_n}, T_n)_{n \geq 0}$$

Analyse de la moyenne et de la distribution de la VAN (t)

Nous avons :

$$VAN(t) = C_t^{(1)} - C_t^{(2)}$$

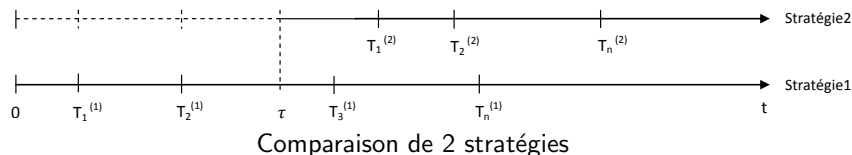
Il s'agit d'évaluer

$$\begin{cases} \mu = \mathbb{E}[VAN(t)] \\ F_{VAN(t)}(u) = \mathbb{P}[VAN(t) \leq u] \end{cases}$$

où u et t sont fixés ($u, t > 0$)

Hypothèse : indépendance des stratégies sachant ce qui s'est passé à l'instant τ

Analyse de la VAN (t)



Pour $t \leq \tau$, nous avons :

$$VAN(t) = 0$$

Pour $t > \tau$, nous avons :

$$\begin{aligned} VAN(t) &= C_t^{(1)} - C_t^{(2)} \\ &= \Gamma \left(\left(Z_v^{(1)}, Z_w^{(2)} \right)_{0 \leq v \leq t, 0 \leq w \leq t - \tau} \right) \end{aligned}$$

Plan

Méthodes de quantification de la $VAN(t)$

Méthode de Monte-Carlo

Méthode de Quasi Monte-Carlo

Comparaison des méthodes

Simulation et estimation par la méthode de Monte-Carlo

Méthode très utilisée pour la simulation des évènements aléatoires

Soit la chaîne de Markov et $i \in \{1, 2\}$:

$$\begin{cases} Z_0^{(i)} \sim f(V^{(i)}) & V^{(i)} \sim \mathcal{U}([0, 1]^m) \\ Z_n^{(i)} = \phi_n(Z_{n-1}^{(i)}, U_n^{(i)}) & n \geq 1 \text{ et } U_n^{(i)} \sim \mathcal{U}([0, 1]) \end{cases}$$

où m représente le nombre de composants du parc

On simule les chaînes de Markov :

$$\begin{aligned} Z_0^{(1)} = f(v^{(1)}) & \xrightarrow{1(Z_0^{(1)}; u_1^{(1)})} Z_1^{(1)} \xrightarrow{2(Z_1^{(1)}; u_2^{(1)})} Z_2^{(1)} \dots Z_{N_t^{(1)}-1}^{(1)} \xrightarrow{N_t^{(1)}(Z_{N_t^{(1)}-1}^{(1)}; u_{N_t^{(1)}}^{(1)})} Z_{N_t^{(1)}}^{(1)} \\ Z_0^{(2)} = f_{Z_{N_{t-\tau}^{(1)}}}^{(1)}(v^{(2)}) & \xrightarrow{1(Z_0^{(2)}; u_1^{(2)})} Z_1^{(2)} \xrightarrow{2(Z_1^{(2)}; u_2^{(2)})} \dots Z_{N_{t-\tau}^{(2)}-1}^{(2)} \xrightarrow{N_{t-\tau}^{(2)}(Z_{N_{t-\tau}^{(2)}-1}^{(2)}; u_{N_{t-\tau}^{(2)}}^{(2)})} Z_{N_{t-\tau}^{(2)}}^{(2)} \end{aligned}$$

On estime la moyenne et la distribution de la VAN (t) par :

$$\hat{\mu} \simeq \frac{1}{N} \sum_{j=1}^N \left(C_{T_{N_t^{(1)}}}^{(1)} - C_{T_{N_{t-\tau}^{(2)}}}^{(2)} \right) \quad \hat{F}_{VAN(t)}(u) \simeq \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{C_{T_{N_t^{(1)}}}^{(1)} - C_{T_{N_{t-\tau}^{(2)}}}^{(2)} \leq u\}}$$

Méthode de Monte-Carlo

▶ Avantages

- + Permet d'estimer facilement tout indicateur
- + Intervalle de confiance des estimations

▶ Inconvénients

- Temps de calcul longs pour des résultats précis

Méthode de Quasi Monte Carlo : principe général

On souhaite estimer :

$$\mathcal{I} = \int_{[0,1]^d} f(u) du \quad (1)$$

Estimation de \mathcal{I} par MC est :

$$\hat{\mathcal{I}} = \frac{1}{N} \sum_{i=1}^N f(u_i), \quad u_i \sim \mathcal{U}([0,1]^d)$$

Idée de la méthode QMC : remplacer l'échantillon de loi uniforme par une suite de points plus uniformément distribués

⇒ **Calcul de la discrédance**

Méthode Quasi Monte-Carlo : la discrédance

Définition (Discrédance extrême)

La discrédance de l'ensemble $P_N = \{u_1, u_2, \dots, u_N\}$ est donnée par :

$$D_N(\mathcal{A}; P_N) = \sup_{A \in \mathcal{A}} \left| \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{(u_i \in A)} - \lambda(A) \right|$$

où $\lambda(A)$ est le volume de A

On a :

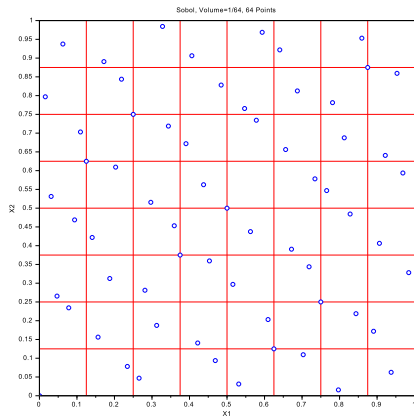
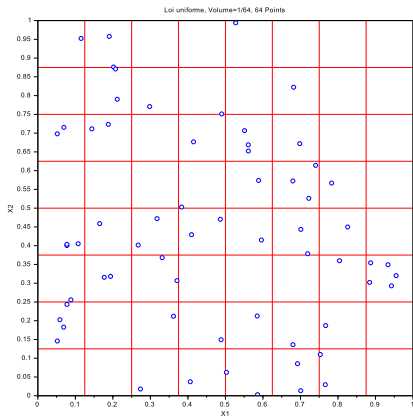
$$0 \leq D_N(\mathcal{A}; P_N) \leq 1$$

\implies **Les suites à faible discrédance**

Exemples de suites à faible discrédance

Noms	Paramètres
Halton	b_1, \dots, b_d
Faure	plus petit $b > d$
Sobol	$b = 2$

Méthode Quasi Monte-Carlo : comparaison



Méthode QMC : évaluation de la loi de la VAN (t)

$$\begin{cases} Z_0^{(i)} \sim f(V^{(i)}) & V^{(i)} \sim \mathcal{U}([0, 1]^m) \\ Z_n^{(i)} = \phi_n(Z_{n-1}^{(i)}, U_n^{(i)}) & n \geq 1 \text{ et } U_n^{(i)} \sim \mathcal{U}([0, 1]) \end{cases}$$

m représente le nombre de composants

$$Z_0^{(1)} = f(V^{(1)}) \xrightarrow{1(Z_0^{(1)}; u_1^{(1)})} Z_1^{(1)} \xrightarrow{2(Z_1^{(1)}; u_2^{(1)})} Z_2^{(1)} \dots \xrightarrow{N_t^{(1)}(Z_{N_t^{(1)}-1}^{(1)}; u_{N_t^{(1)}}^{(1)})} Z_{N_t^{(1)}}^{(1)}$$

$$Z_0^{(2)} = f_{Z_{N_{t-\tau}^{(1)}}} (V^{(2)}) \xrightarrow{1(Z_0^{(2)}; u_1^{(2)})} Z_1^{(2)} \xrightarrow{2(Z_1^{(2)}; u_2^{(2)})} \dots \xrightarrow{N_{t-\tau}^{(2)}(Z_{N_{t-\tau}^{(2)}-1}^{(2)}; u_{N_{t-\tau}^{(2)}}^{(2)})} Z_{N_{t-\tau}^{(2)}}^{(2)}$$

On doit construire une suite à faible discrédance (SFD) de dimension

$$N_t^{(1)} + N_{t-\tau}^{(2)} + 2m$$

Méthode QMC : évaluation de la loi de la VAN (t)

On choisit 2 entiers $n^{(1)}, n^{(2)} \in \mathbb{N}^*$ tels que

$\mathbb{P} \left(n^{(1)} > N_t^{(1)} + m \right) > 1 - \varepsilon$ et $\mathbb{P} \left(n^{(2)} > N_{t-}^{(2)} + m \right) > 1 - \varepsilon$, pour $\varepsilon > 0$

On construit une SFD S de dimension $n^{(1)} + n^{(2)}$ et on utilise les $n^{(1)}$ premières composantes pour la stratégie 1 et les $n^{(2)}$ dernières pour la stratégie 2

Si on pose $S^{(1)} = S(1 : n^{(1)})$ et $S^{(2)} = S(n^{(1)} + 1 : n^{(1)} + n^{(2)})$, alors

$$Z_0^{(1)} = f(S^{(1)}(1 : m)) \quad \begin{matrix} 1 \\ \longrightarrow \end{matrix} \begin{matrix} (Z_0^{(1)}; S^{(1)}(m+1)) \\ \longrightarrow \end{matrix} Z_1^{(1)} \quad \begin{matrix} 2 \\ \longrightarrow \end{matrix} \begin{matrix} (Z_1^{(1)}; S^{(1)}(m+2)) \\ \longrightarrow \end{matrix} Z_2^{(1)} \quad \dots \quad Z_{N_t^{(1)}}^{(1)}$$

$$Z_0^{(2)} = f_{Z_{N_{t-}^{(2)}}^{(1)}}(S^{(2)}(1 : m)) \quad \begin{matrix} 1 \\ \longrightarrow \end{matrix} \begin{matrix} (Z_0^{(2)}; S^{(2)}(m+1)) \\ \longrightarrow \end{matrix} Z_1^{(2)} \quad \begin{matrix} 2 \\ \longrightarrow \end{matrix} \begin{matrix} (Z_1^{(2)}; S^{(2)}(m+2)) \\ \longrightarrow \end{matrix} Z_2^{(2)} \quad \dots \quad Z_{N_{t-}^{(2)}}^{(2)}$$

On évalue ainsi les trajectoires de la VAN (t) jusqu'à l'horizon de temps t fixé et on en déduit l'estimation de la moyenne et de loi de la VAN (t) par la loi des grands nombres

Méthode Array Quasi Monte Carlo (L'Ecuyer, Lécot, Tuffin)

Idée :

Réduire la dimension de le SFD

Principe :

- ▶ Simuler plusieurs chaînes de Markov en parallèle
- ▶ Trier par ordre croissant les chaînes pour améliorer la distribution de la chaîne de Markov à chaque saut

Méthode AQMC : évaluation de la loi de la VAN (t)

On a :

$$\begin{cases} Z_0^{(i)} \sim f(V^{(i)}) & V^{(i)} \sim \mathcal{U}([0, 1]^m) \\ Z_n^{(i)} = \phi_n(Z_{n-1}^{(i)}, U_n^{(i)}) & n \geq 1 \text{ et } U_n^{(i)} \sim \mathcal{U}([0, 1]) \end{cases}$$

- ▶ on considère N termes de deux SFD de dimension m et 1
- ▶ on construit N CM en parallèle
- ▶ on trie les CM par rapport aux instants de sauts des chaînes

On a alors :

$$\begin{array}{ccccccc} Z_{0,1}^{(1)} = f(v_1) & \phi_1 \left(\begin{array}{c} Z_{0,1}^{(1)} \\ \longrightarrow \\ u_{11} \end{array} \right) & Z_{1,1}^{(1)} & \xrightarrow{\text{tri}} & Z_{1,1}^{(1)} := Z_{1,(1)}^{(1)} & \phi_2 \left(\begin{array}{c} Z_{1,1}^{(1)} \\ \longrightarrow \\ u_{N+1} \end{array} \right) & Z_{2,1}^{(1)} \dots \\ Z_{0,2}^{(1)} = f(v_2) & \phi_1 \left(\begin{array}{c} Z_{0,2}^{(1)} \\ \longrightarrow \\ u_{12} \end{array} \right) & Z_{1,2}^{(1)} & \longrightarrow & Z_{1,2}^{(1)} := Z_{1,(2)}^{(1)} & \phi_2 \left(\begin{array}{c} Z_{1,2}^{(1)} \\ \longrightarrow \\ u_{N+2} \end{array} \right) & Z_{2,2}^{(1)} \dots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ Z_{0,N}^{(1)} = f(v_N) & \phi_1 \left(\begin{array}{c} Z_{0,N}^{(1)} \\ \longrightarrow \\ u_{1N} \end{array} \right) & Z_{1,N}^{(1)} & \longrightarrow & Z_{1,N}^{(1)} := Z_{1,(N)}^{(1)} & \phi_2 \left(\begin{array}{c} Z_{1,N}^{(1)} \\ \longrightarrow \\ u_{2N} \end{array} \right) & Z_{2,N}^{(1)} \dots \end{array}$$

Méthode QMC : limites et solution

Limites

- ▶ Pas d'intervalle de confiance pour les estimations
- ▶ Une borne supérieure de l'erreur difficile à calculer

Solution : randomisation des SFD

- ▶ La méthode Randomized Quasi Monte Carlo (RQMC)
- ▶ La méthode Randomized Array Quasi Monte Carlo (RAQMC)

Comparaison des méthodes

Comparaison des méthodes déterministes (QMC et AQMC) via les indicateurs suivants :

- ▶ l'erreur relative sur la $VAN(t)$ moyenne :

$$\varepsilon = \left| \frac{\hat{\mu} - \mu}{\mu} \right|$$

où $\hat{\mu} = \frac{1}{N} \sum_{i=1}^N VAN_i(t)$ et μ la $VAN(t)$ moyenne de référence,

- ▶ l'erreur relative sur la loi de la $VAN(t)$: on considère un ensemble $I = \{a_1, \dots, a_{n_I}\}$ et on définit :

$$\mathcal{E} = \frac{1}{n_I} \sum_{j=1}^{n_I} \epsilon_{a_j}$$

où $\epsilon_{a_j} = \left| \frac{\hat{P}_{VAN(t)}^{a_j} - P_{VAN(t)}^{a_j, ref}}{P_{VAN(t)}^{a_j, ref}} \right|$, $\hat{P}_{VAN(t)}^{a_j} = \frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{VAN_i(t) \leq a_j\}}$ et $P_{VAN(t)}^{a_j, ref}$ la valeur de référence

Comparaison des méthodes

Comparaison des méthodes aléatoires (MC, RQMC et RAQMC) via les indicateurs suivants :

- ▶ l'erreur relative moyenne sur la $VAN(t)$ moyenne définie par :

$$\bar{\varepsilon} = \frac{1}{J} \sum_{j=1}^J \varepsilon_j$$

$\varepsilon_j = \left| \frac{\hat{\mu}_j - \mu}{\mu} \right|$ est l'erreur relative sur la $j^{i\grave{e}me}$ simulation de la méthode MC ou randomisation de la méthode (A)QMC

- ▶ le facteur de réduction de variance VRF défini par :

$$VRF = \frac{\mathbb{V}_{R(A)QMC}(\hat{\mu})}{\mathbb{V}_{MC}(\hat{\mu})}$$

Plan

Résultats numériques

Présentation du problème

On considère 2 composants identiques et indépendants et un stock commun de taille n_0

Stratégies à comparer :

- ▶ Stratégie 1 (stratégie corrective) : si une pièce de rechange est disponible alors remise à neuf du composant en panne, sinon le composant devient indisponible
- ▶ Stratégie 2 (stratégie préventive) : remplacement de tous les composants à l'instant τ + remplacements correctifs

Résultats numériques

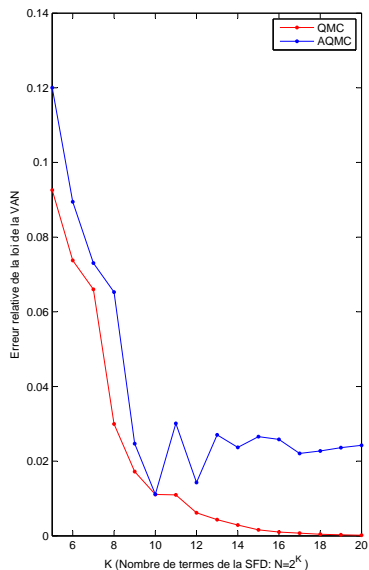
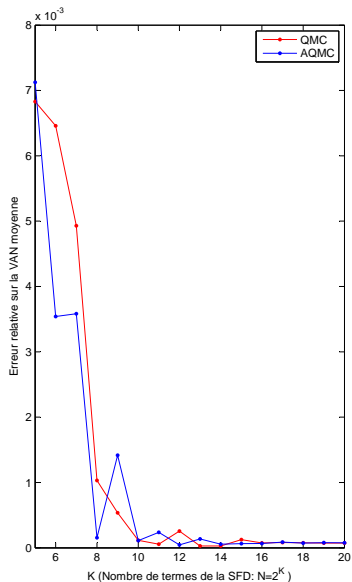
Logistique du stock :

- ▶ Si le stock est vide alors, une nouvelle pièce est commandée
- ▶ Si un composant tombe en panne alors qu'il y a pas de pièce de rechange et qu'une commande est déjà en cours, une pièce supplémentaire est commandée
- ▶ Annulation de toutes les commandes en cours et des coûts induits sauf la première commande à l'instant τ
- ▶ Délai d'approvisionnement déterministe ν

Objectif : évaluer la moyenne et la distribution de la VAN à 60 ans

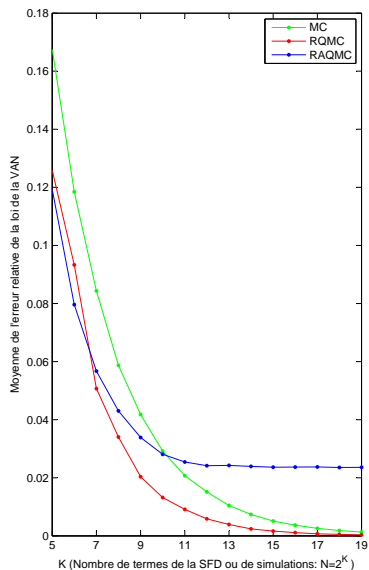
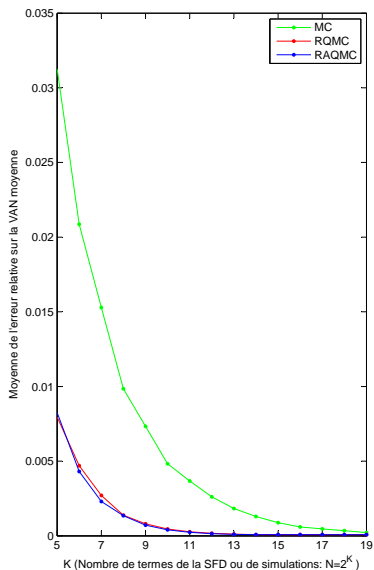
Résultats

Méthodes déterministes : QMC et AQMC



Résultats

Méthodes aléatoires : MC, RQMC et RAQMC



Résultats numériques

Méthodes aléatoires : MC, RQMC et RAQMC

K	11	14	17
VRF_{RQMC}	0.0051	0.0016	0.0005
VRF_{RAQMC}	0.0031	0.0011	0.0003

Facteur de réduction de variance sur la $VAN(60)$ moyenne avec les $J = 500$ simulations et $N = 2^K$

$a \backslash K$	11		14		17	
	RQMC	RAQMC	RQMC	RAQMC	RQMC	RAQMC
-250 (0.4118)	0.1080	0.2041	0.0593	0.2323	0.0308	0.1926
-200 (0.7877)	0.0960	0.1188	0.0677	0.1550	0.0367	0.1483
-150 (0.9267)	0.1160	0.1835	0.1056	0.1853	0.0595	0.1976

Facteur de réduction de variance sur $\mathbb{P}(VAN(60) \leq a)$ avec $J = 500$ simulations et $N = 2^K$

Conclusion & Perspectives

- ▶ Méthode QMC
 - ▶ Bonne estimation de la distribution de la $VAN(t)$
 - ▶ Choix de la dimension de la SFD
 - ▶ Dégradation de la méthode lorsqu'on augmente le nombre de composants
- ▶ Méthode AQMC
 - ▶ Bonne estimation de la $VAN(t)$ moyenne
 - ▶ Erreur commise sur la loi de la $VAN(t)$ (diminution de l'erreur avec augmentation du nombre de composants)
 - ▶ Choix de la fonction de tri des chaînes
- ▶ Couplage des méthodes à un algorithme d'optimisation (algorithme génétique)

MERCI DE VOTRE ATTENTION



Méthode AQMC : justification du tri

Exemple :

Soit

$$\begin{cases} X_0 = x_0 \\ X_n = \phi_n(X_{n-1}, U_n) \end{cases} \quad n \geq 1 \text{ et } U_n \sim \mathcal{U}([0, 1])$$

On souhaite estimer la loi X_n

Méthode AQMC : justification du tri

Soit $x \in \mathbb{R}$, on a :

$$\begin{aligned} F_n(x) &= \mathbb{P}[X_n \leq x] \\ &= \mathbb{P}[\phi_n(X_{n-1}, U_n) \leq x] \\ &= \mathbb{P}[\phi_n(F_{n-1}^{-1}(V), U_n) \leq x] \\ &\approx \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{n(F_{n-1}^{-1}(v_j); u_j) \leq x\}} \\ &\approx \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{n(F_{n-1}^{-1}(\frac{j-0.5}{N}); u_j) \leq x\}} \\ &\approx \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{n(q_{n-1}(\frac{j-0.5}{N}); u_j) \leq x\}} \\ &\approx \frac{1}{N} \sum_{j=1}^N \mathbf{1}_{\{n(x_{n-1}^{(j)}; u_j) \leq x\}} \end{aligned}$$

Résultats numériques

Méthodes déterministes : QMC et AQMC

Méthodes	$\mathbb{E} [VAN(60)]$	\mathcal{E}
Référence	-231.0231 [-231.0527, -230.9935]	-
QMC	-231.0299	0.0029
AQMC	-231.0362	0.0237
MC	-231.2346 [-231.9670, -230.5022]	0.0081

Estimation $VAN(60)$ moyenne et erreur relative sur la loi de la $VAN(60)$ avec $N = 2^{14} = 16384$

Résultats numériques

Méthodes déterministes : QMC et AQMC

a	-250	-200	-150
Référence	0.4118 [0.4115, 0.4121]	0.7877 [0.7874, 0.7879]	0.9267 [0.9266, 0.9269]

Méthodes \ a	-250	-200	-150
QMC	0.4121	0.7879	0.9267
AQMC	0.4183	0.7893	0.9273
MC	0.4101 [0.4074, 0.4127]	0.7859 [0.7837, 0.7882]	0.9263 [0.9249, 0.9277]

Estimation de la fonction de répartition en $\{-250, -200, -150\}$ avec $N = 2^{17} = 131072$